

## 映画の映像理論に基づく対面会議シーンの自動撮影手法

井上 亮 文<sup>†1</sup> 吉田 竜 二<sup>†2</sup> 平石 絢 子<sup>†3</sup>  
重野 寛<sup>†4</sup> 岡田 謙 一<sup>†4</sup> 松下 温<sup>†5</sup>

本研究では、映画の映像理論に基づく対面会議シーンの自動撮影方法を提案する。対面会議シーンの撮影には複数のカメラを用いる必要があるが、位置的に離れたカメラの映像でスイッチングを行うと急激な変化が生じ、視聴者が参加者の位置関係の認識に混乱を来す。提案手法では、人物の位置関係を明確にする映像理論であるイマジナリーラインを設定および解除する方法と、参加者の対話シーンを強調する三角形配置に従った撮影用カメラの決定方法を定義している。実験では正確なイマジナリーラインを冗長な部分があるものの70%の確率で設定でき、映像表現に影響を与えるような検出口は少なかった。アンケートによる比較実験では、人手でスイッチングしたものと比べて位置関係の見やすい映像を自動生成できたことが確認された。

### Automatic Shooting Method for Face-to-face Meetings Based on Grammar of the Film Language

AKIFUMI INOUE,<sup>†1</sup> RYUJI YOSHIDA,<sup>†2</sup> JUNKO HIRAISHI,<sup>†3</sup>  
HIROSHI SHIGENO,<sup>†4</sup> KEN-ICHI OKADA<sup>†4</sup> and YUTAKA MATSUSHITA<sup>†5</sup>

In this paper, an automatic shooting method for face-to-face meeting scene based on grammar of the film language is proposed. Shooting a meeting scene requires multiple video cameras, however, viewers may get confused in case of switching shot between spatially distant cameras. To shoot participants effectively, we introduce two filming theory into automatic shooting method. The detection of the "imaginary line" makes spatial relationships of the participants clear. The determination of the "camera triangle" figures out the conversation of the participants. Although there remains some redundancies, the detection ratio of the ideal imaginary line was 70%, the loss of which had less effects on the video image. The comparative experimental result indicated the availability.

#### 1. はじめに

映像コンテンツの制作を効率的に行うため、様々なシーンの自動撮影技術が注目されている<sup>1)~3)</sup>。なかでも対面会議シーンの自動撮影は会議がどのように進化したかを明確に伝えるマルチメディア議事録として利用可能である。従来の会議シーンは固定カメラや全方位カメラ<sup>4),5)</sup>で撮影されるなど単調であり、映像を視聴者により効果的に提示するためには何らかの演出

(映像表現)を加える必要があると考えられる。

これまで会議に関連した映像表現は、主として遠隔会議システムでの利用を想定してきた。テレビ番組のカメラワークの知識に基づいた多人数遠隔会議システムでは、1地点の参加者を様々なショットで撮影し、それらを適切なタイミングで切り替えた(スイッチングした)映像を遠隔の参加者に提示することで、視聴者が退屈しない映像を作り出すことに成功している<sup>6),7)</sup>。しかし、対面会議シーンの映像化では、退屈させない映像のスピード感よりも、会議の様子や内容の確実な伝達が重視されると考えられる。

また、遠隔会議システムにおける撮影環境は、1列に配置された参加者をその前方に設置した1台のカメラで撮影するものである。これに対し対面会議は、参加者が机を囲む円卓型の座席配置で行うことがほとんどである。カメラが空間的制限なく自由に移動可能な仮想空間<sup>8),9)</sup>とは異なり、実空間の対面会議シーンの全参加者を確実に撮影するには1台のカメラでは困難

†1 慶應義塾大学大学院理工学研究科  
Graduate School of Science and Technology, Keio University

†2 日本ヒューレット・パッカド株式会社  
Hewlett-Packard Japan, Ltd.

†3 株式会社エヌ・ティ・ティ・ドコモ  
NTT DoCoMo, Inc.

†4 慶應義塾大学理工学部  
Faculty of Science and Technology, Keio University

†5 東京工科大学  
Tokyo University of Technology

であり、複数台のカメラをあらかじめ固定した位置に設置することになる。しかし、設置位置が大幅に異なるカメラの映像でスイッチングをすると、視点が急激に変化して参加者間の位置関係や方向感覚を失い、視聴者がどこからどの対象を見ているかを認識できなくなってしまう<sup>10)</sup>。

このように対面会議の撮影は、遠隔会議のそれと比べて、必要な映像表現と撮影環境が大きく異なる。蓄積映像の編集<sup>11)</sup>には大きな時間と労力を要するうえ、適切な映像を選択するには専門的な知識も必要となる。複数カメラを用いた撮影で会議内容を効果的に伝達する効果的な映像を、専門のオペレータに頼ることなくリアルタイムで取得する手法が必要である。

そこで本論文では、映画の映像理論に基づく対面会議シーンの効果的な自動撮影手法を提案する<sup>12)</sup>。提案手法では、人物の位置関係を明確にするイマジナリーラインとカメラの三角形配置という2つの映像理論に注目した。この映像理論を対面会議シーンに適用し、参加者の発言の推移からイマジナリーラインを検出してシーンを効果的に撮影するカメラを自動的に決定する。プロトタイプシステムを実装して評価を行った結果、本手法の有効性を確認した。

以下、2章では本研究で用いる映像理論について、3章では提案手法について、4章ではプロトタイプの実装について述べる。5章で実験と結果に対する考察を述べ、6章をまとめおよび今後の課題とする。

## 2. 映画の映像理論

### 2.1 イマジナリーラインと三角配置

シーンの中心となる2人の人物には、2人の間に交わされる目線の方向に基づいた、相手に関心を示す直線が流れている。この直線はイマジナリーラインと呼ばれる<sup>13)</sup>。そしてこのイマジナリーラインを底辺として三角形を形成する位置にカメラを配置するのがカメラの三角形配置である(図1)。

図1におけるカメラ1から3では、人物Aを画面の左側でとらえるのに対し、カメラ4では反対の右側でとらえる(図2)。ここでカメラ1, 2, 3は、イマジナリーラインで分割された空間においてすべて同じ側に属している。このようにイマジナリーラインを越えないカメラの映像でスイッチングを行うことで、各人物の位置関係を明確にすることが可能となる。

また三角形配置されたカメラのうち、イマジナリーラインを底辺とするカメラ2台(底辺カメラ)からは、2人のうちどちらか一方を強調したショットが得られる。たとえばカメラ2はカメラ1より人物Aの表

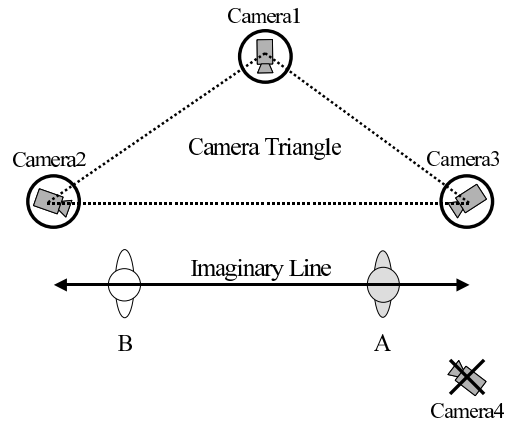


図1 イマジナリーラインとカメラの三角形配置  
Fig. 1 Imaginary line and camera triangle.

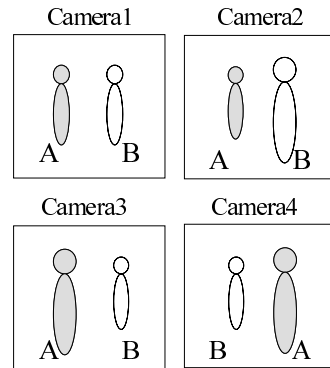


図2 各カメラの視点  
Fig. 2 Each camera's view.

情をよりよく映せる。一方頂角に位置するカメラ(頂角カメラ)からは、2人を同じように映したマスターショット的な映像が得られる。この3つのショットを切り替えるとイマジナリーライン上にいる人物を強調できる。

対面会議シーンを自動撮影する場合にこのイマジナリーラインをカメラの決定基準として考慮することで、設置位置の異なるカメラの映像を切り替えた際に生じる位置・方向間隔の混乱を防ぐことができる。そして三角形配置されたカメラの映像を用いてスイッチングを行うことで参加者間の対話を強調し、違和感の少ない映像を生成できると考えられる。

### 2.2 映画の分析

人物を撮影する際、イマジナリーラインは必ずしも越えてはいけないものではない。あえてこれを越えることで場面転換や時間経過を表現し、映像がより興味深く演出される場合もある<sup>14)</sup>。そこで、実際の対面会議シーンの映像におけるイマジナリーラインの扱いを確認するための分析を行った。その対象として「12人

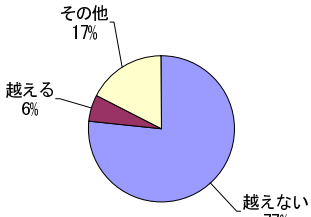


図3 スイッチングとイマジナリーラインの関係

Fig. 3 The relationship between an imaginary line and switching.

の怒れる男」という映画を取り上げた。この作品は12人の陪審員が殺人事件について審理を続けるというストーリーであり、ほぼ1つの部屋の対面会議シーンだけで作品全体117分を構成している。よって分析対象として最適であると考えた。調査したのは、作品の中で2人の人物の対話シーンにおけるスイッチング253回である。

分析の結果を図3に示す。全体の77%である194回はイマジナリーラインを越えないことが分かった。越える場合も、「他の対象を撮影した、方向感覚を感じさせないショットをはさむ」「イマジナリーライン上の映像をはさむ」といった違和感を与えない中間ショットを挿入する配慮がなされていた。残りの44回は、まったく別の演出ショットへの切替えとその復帰であった。

分析結果より、対面会議シーンを複数カメラで撮影する際には、あえてイマジナリーラインを越える演出よりも、参加者の位置関係を重視した演出をすべきであるといえる。

### 3. 提案方式

本章では、対面会議シーンに対してイマジナリーラインとカメラの三角形配置とを組み込んだ自動撮影手法について述べる。ここで、対面会議の状況説明に用いる用語について説明しておく。1人の参加者が意見を述べることを「発言」、一定時間特定の参加者2人が相互に発言を繰り返すことを「対話」、3人以上の参加者が次々と発言をすることを「議論」とする。

#### 3.1 イマジナリーラインの定義

本研究では10人以下の比較的小規模の対面会議を対象とし、各参加者はあらかじめ決められた座席位置に着席したまま会議を進める。この状況において対話・議論をしている参加者の位置を特定し、その間に存在する線分をイマジナリーラインとして定義する。

イマジナリーラインはシーンの中心となる人物の間に存在するが、この「シーンの中心」としては、発言を繰り返して直接対話をする2人、次々と発言をして議

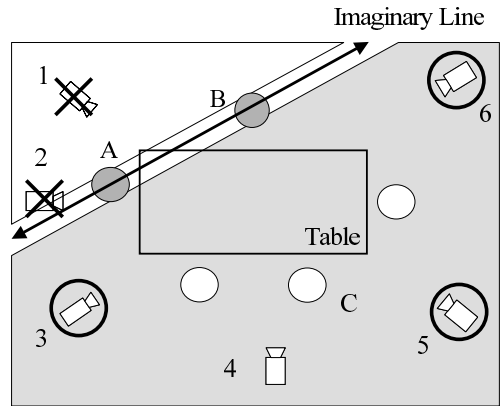


図4 2人の対話におけるイマジナリーラインの設定

Fig. 4 Establishment between two participants.

論をする3人、の2種類を対象とする。逆に、シーンの中心となる人物が存在しない場合にはイマジナリーラインも存在しない。これには多くの参加者(4人以上)が次々と発言して、中心的な役割が参加者の間で拡散する場合が相当する。

このように本研究では3人までの議論に対してイマジナリーラインをベースにした撮影を行うが、つねに同じ人物が対話・議論をし、その位置関係が変化しないのであれば固定的なイマジナリーラインを想定できる。しかし、複数の参加者が互いに議論するある程度の長さの会議では、発言者とその直接的な相手はつねに変化するため、イマジナリーラインも時々刻々と変化する。よってこれをいつ設定、解除するのかを定義する必要がある。

#### 3.2 対話におけるイマジナリーラインの設定

提案手法では各参加者がマイクを持ち、その発言を音声情報として検出することで現在の発言者を判断する。そして、たとえば  $A \rightarrow B \rightarrow A$  のようにある2人の参加者の間で交互に発言が繰り返された場合を「対話」が発生したと見なし、AとBの間にイマジナリーラインを設定する(図4)。

#### 3.3 議論におけるイマジナリーラインの設定

実際の会議では、1人が進行役として他の2人の発言をまとめる3人での議論もしばしば存在する。3.2節の手法は各発言者の間にイマジナリーラインを設定するため、A、B、C3人の議論にこれをそのまま適用するとAB、BC、BAの間に計3本が設定されることになる。この3本のイマジナリーラインをいずれも越えないような位置に設置されたカメラは、その条件の厳しさから存在しない可能性がある。

そこで提案手法では、3人のうち2人をグループ化して1対2での対話と見なすことで、3本のイマジナリー

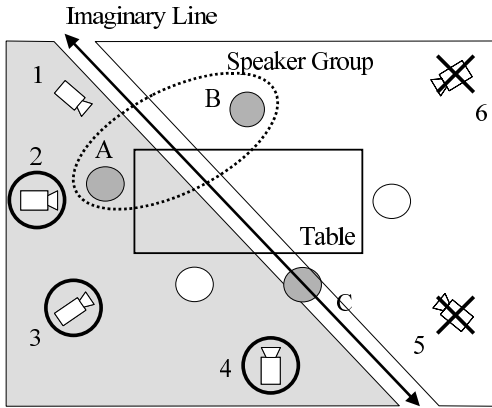


図 5 3人の議論におけるイマジナリーラインの設定  
Fig. 5 Establishment among three participants.

ラインを1本に簡略化する．図5にその例を示す．AとBの対話によりイマジナリーラインが設定されている状況で，Cが発言をし，その後再びAかBのどちらかが発言をしたとする．この流れは，ABのうちどちらか一方の発言を(A|B)と表すと(A|B) → C → (A|B)となる．これは，ABのグループとCの間で3.2節の手法が適用されることに等しい．そこでAとBの間に設定されたイマジナリーラインを，AとBの中間の座標とCの座標との間に新たに設定しなおす．

3.4 イマジナリーラインの解除

対話，議論のいずれも場合でも，設定されたイマジナリーラインと関係のない参加者の発言が2回続いたところでこれを解除する．ここで2回としたのは，1回では3人の議論に対応できないことと，3回以上では条件があまりに厳しく，設定されたイマジナリーラインがなかなか解除されないためである．

3.5 撮影カメラの決定

シーンの中心的な役割を果たす対話・議論中の参加者を撮影するカメラ(撮影カメラ)は，設定されたイマジナリーラインを越えない三角形配置をとるように決定する．

イマジナリーラインが設定されている場合は，その直線によって会議空間が2分割できる．そこでまずそれぞれの空間に存在し，直線上の2人を撮影可能なカメラの台数をカウントし，台数が多い方の空間からカメラを選択する．この際，頂角カメラはイマジナリーラインから遠く，2人を概観できるものにする．逆に底辺カメラ2台は，イマジナリーラインに近くなるよう決定する．例として図4ではカメラ3, 5, 6を，図5ではカメラ2, 3, 4を撮影カメラとして決定さ

表 1 ショットの分類  
Table 1 The shot classification.

ショット	説明
全景	参加者全員を映す
発言者	発言者とその周囲の参加者を映す
演出 1	底辺カメラから対話・議論中の参加者を映す
演出 2	頂角カメラから対話・議論中の参加者を映す

表 2 1 ショットの持続時間と出現確率  
Table 2 The duration probability.

持続時間(秒)	出現確率
2.5	55%
7.5	30%
12.5	10%
17.5	5%

れる．

このように三角形配置のカメラを求めることで，それぞれのカメラからは印象がまったく異なるショットが得られる．これらをスイッチングすることで，映像にメリハリのある効果的なカメラワークが実現できると考えられる．

3.6 スイッチング

退屈な映像にならないよう，各カメラから得られるショットをスイッチングして1本の映像ストリームを生成する．スイッチングに関しては基本的に従来の知見を利用する．

スイッチングに用いる4種類のショットを表1に示す．会議が行われている最中は，会議空間全体を映すショット(全景)と，発言者を前方から写すショット(発言者)がつねに用意されている．イマジナリーラインが設定された場合は，そのラインに関わる参加者を演出するため，新たに底辺カメラ2台のショット(演出1)と頂角カメラからのショット(演出2)が用意される．

1ショットの持続時間は，発言者が交代したときと，同一ショットが長時間続き単調な映像になるのを避けるときの2パターンがある．特に後者は，持続時間の長いショットほどその出現回数が減少することが分かっている<sup>6)</sup>．本研究もこれに習い，2.2節の映画を分析から，ショットの持続時間とその出現確率の関係を表2のように設定した．

これら表1と表2のデータを用いて会議の進行と並行してスイッチングを行う．シーンの中心となる参加者が存在せずイマジナリーラインが設定されない場合は，全景ショットと発言者ショットの間でスイッチングをする．対話・議論が発生し中心的な参加者がいる場合は，それら参加者たちの位置関係を明確にする

もしくは1人+2人であるが，簡単のため2人と表記する．

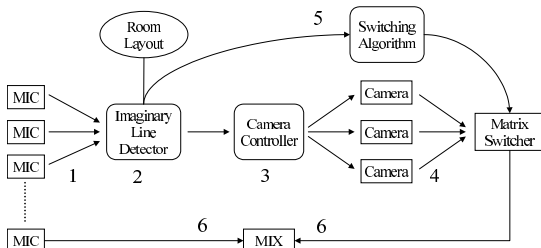


図 6 プロトタイプによる撮影の流れ

Fig. 6 Shooting processes of the prototype system.

ために、底辺カメラと頂角カメラによる 3 つの演出ショットでスイッチングを行う。

#### 4. 実 装

提案手法に基づいて、対面会議シーンを自動撮影するプロトタイプシステムを構築した。プロトタイプにおける撮影の流れは以下のステップからなる(図 6)。

- (1) イマジナリーラインの設定に必要な発言者の特定は、参加者がマイクに向かって話すことでシステムに認識させる。0.5 秒間に 4000 回のサンプリングを行い、閾値以上の入力があった時点で発言者と判定する。
- (2) (1) で検出した発言者と会議空間のレイアウトをもとに、イマジナリーラインの設定・解除を行う。そしてその有無に応じて撮影カメラを決定する。
- (3) イマジナリーラインが設定されている場合、撮影カメラは該当する参加者を映すようレンズの方向などが制御される。
- (4) それぞれのカメラの映像出力はマトリックススイッチャーへ入力されている。このスイッチャーを制御することにより、任意の 1 つを出力映像として選択できる。
- (5) スwitching アルゴリズムに基づいてスイッチャーを適切に制御する。
- (6) 各カメラの入力映像が 1 本の映像ストリームとして出力される。この映像は会議空間全体の音声とミックスされる。

##### 4.1 実装環境

イマジナリーライン検出部およびカメラ選択・制御部の実装には Windows2000, Pentium III 600 MHz の PC を使用した。システムの各モジュールはすべて Java 言語で実装した。使用したカメラは Canon 社製 VC-CI である。この雲台付きカメラは RS-232C シリアルポート経由で PC から制御可能であり、首振り位置のプリセットメモリ機能を保持している。マトリッ

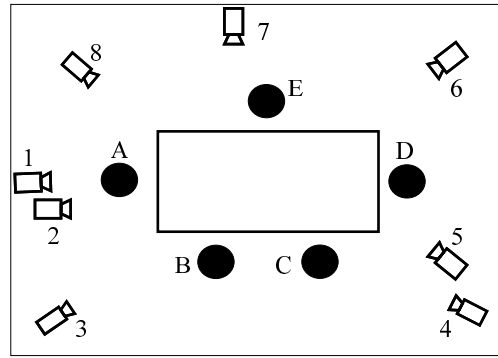


図 7 会議空間のレイアウト

Fig. 7 Meeting space layout.

表 3 各カメラのショット

Table 3 Each camera's shot.

Camera	shot 1	shot 2	shot 3
1	ABCDE	-	-
2	BCD	CDE	-
3	CDE	AE	-
4	ABE	-	-
5	ADE	ACE	-
6	ABC	BCD	ABE
7	BC	-	-
8	ABC	BCD	CDE

クススイッチャーは IMAGENICS 社製 SW-1010F を用いた。

##### 4.2 システム構成

プロトタイプの構成を図 7 に示す。参加者は全員着席しているものの、会議中は着席位置に多少の移動が見られる。今回実験に用いたカメラには被写体位置補正機能がないため、ズームを引き気味にして画面内に複数の参加者が映るショットになるようにした。

また、パンやチルトといったカメラの動作途中の映像が頻繁に登場すると、視聴者は映像に強い違和感を覚えてしまい、本研究の目指す評価が正しく得られない可能性がある。そこで、できるだけこのような動作をしないように、参加者人数に対してカメラの設置位置を 8 カ所と多めにした。各カメラのショットを表 3 に、2 者間でイマジナリーラインが設定された場合の撮影カメラを表 4 に示す。カメラ 1 が会議空間全体の撮影専用なのを除いて、他のカメラは 1~3 通りのショットを保持し、常時 2 人もしくは 3 人を画面内におさめている。例として BC の対話時のショットを図 8 に示す。イマジナリーラインの位置に基づきカメラ 2, 6, 7 が選択されている。

なお、最も離れた A と D の対話時には、撮影空間の広さの限界から 1 台しか選択できなかった。またシ

表 4 2 者間対話における撮影カメラ

Table 4 Triangle cameras at two participants' dialogue.

参加者 \ 相手	B	C	D	E
A	4,6,8	5,6,8	5	3,4,6
B	-	2,6,7	2,6,8	4,6
C	-	-	3,6,8	2,3,8
D	-	-	-	3,5,8

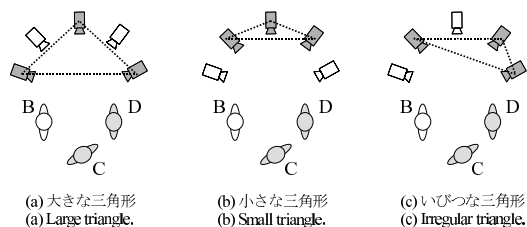


図 9 比較されるカメラ配置

Fig.9 Compared camera triangles.

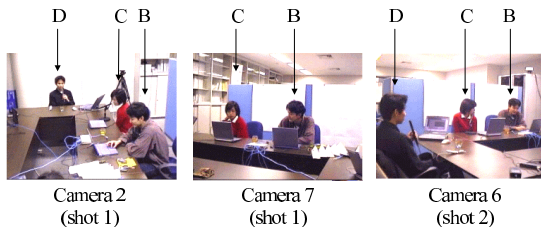


図 8 プロトタイプにおけるスイッチング例

Fig. 8 An example of edited image.

システムがイマジナリーラインを認識する前にスイッチングが起こる場合も見られた。これら実装上の問題は今後の課題とする。

### 5. 評価実験

#### 5.1 イマジナリーライン検出方法の評価

提案手法でどの程度イマジナリーラインを検出できたかを評価するため、図 7 のレイアウトによる対面会議を 10 分間撮影した。このとき、表 3 のすべてのショットを用意しておき、撮影後に会話内容の分析と参加者の顔の向きから理想的なイマジナリーラインの発生時刻・位置を 1 秒ごとに手動で決定し、提案手法で検出したイマジナリーラインの位置・時刻と比較を行った。また、両者のショットをそれぞれのイマジナリーラインをもとに編集し、映像表現に与える影響についても分析した。

#### 5.2 撮影カメラ決定方法の評価

撮影カメラ決定方法の違いによる影響を調べるため、図 7 の B, C, D による対話・議論を 3 つのカメラ配置で撮影した。図 9 にこの配置を示す。(a) は提案手法の配置であり、人物から頂角カメラを遠く、底辺カメラを近くに配置する。(b) は底辺カメラを遠ざけ、より小さな三角形を構成する。(c) は他の 2 つに比べていびつな配置である。やはりスイッチングのタイミングは 3 つの映像ですべて共通とした。

次にそれぞれの映像(約 5 分)を大学生の被験者 16 人に見比べてもらい、1 分ごとにどの映像が好ましいかを順位付けするように指示した。1 位に選ばれていた場合には 3 点、2 位の場合は 2 点、3 位ならば 1 点

として点数をつけ、順位付けの理由も簡潔に記述してもらった。

#### 5.3 アンケート評価

本研究で意図した効果が実際に映像に現れているかを確認するため、プロトタイプで自動撮影した映像を大学生の被験者 16 人に見てもらい、アンケートに 5 段階で評価してもらった。比較評価用に作成した映像は、特別な撮影知識を持たないカメラマン 1 人が手動で撮影カメラとショット持続時間を決定したものである。

アンケートの質問項目を表 5 に示す。イマジナリーラインの効果を比較するため、映像から伝わる位置関係に関する質問項目を用意した(項目 4, 5, 6)。また、カメラの三角形配置の効果を比較するために、参加者の映り具合に関する質問を用意した(項目 1, 2, 8, 9, 11)。残りの項目は映像演出用のスイッチングに関するものである(項目 3, 7, 10, 12)。

また、アンケート以外にも気になった点やシステムへの要求など自由なコメントを記入してもらった。

#### 5.4 結果および考察

##### 5.4.1 検出精度の影響

結果をタイムチャートとして図 10 に示す。濃色の部分はイマジナリーラインが設定されていた時間帯を表しており、設定場所については考慮していない。右横の括弧内の数値は映像全体においてイマジナリーラインが設定されていた時間の占める割合である。

イマジナリーラインを正確な位置・時刻に設定できた割合を表すカバー率  $P$  は、手動による設定時間の合計  $T_i$  と、提案手法による位置・時刻が手動設定のそれと一致した時間の合計  $T_c$  を用いて次のように定義する。

$$P = \frac{T_c}{T_i} \times 100 (\%) \tag{1}$$

提案手法の設定時間の合計は手動よりも約 50 秒長く、冗長な設定が含まれている。そこで、いかに無駄なく正確な位置・時刻にイマジナリーラインを設定できたかを有効率  $E$  として定義する。この  $E$  は、提案

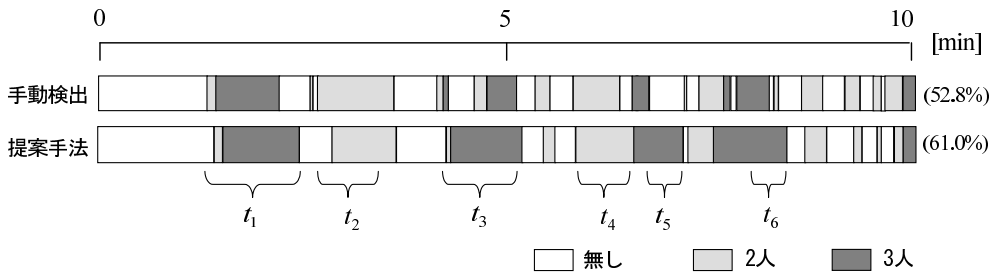


図 10 イマジナリーライン検出のタイムチャート

Fig. 10 The time chart of the imaginary line detection.

手法による設定時間の合計  $T_m$  と、先ほどの  $T_c$  を用いて次のように表すことができる。

$$E = \frac{T_c}{T_m} \times 100(\%) \quad (2)$$

この実験でのカバー率  $P$  は 70%であった。内訳は、2人の対話時が 57%、3人の議論時が 93%となった。一方、有効率  $E$  は 61%であり、同様に内訳は 2人の場合に 67%、3人の場合は 56%となった。結果より、2人の場合はカバー率は低いが有効率は 3人に比べて良い。逆に 3人の場合はカバー率は高いが有効率が低く、正確な位置と時間帯に設定はできたものの、それと同じくらい unnecessary な設定も存在していた。

この原因は設定と解除にかかる遅延である。図 10 中  $t_1$  の前半部分では、B が D に向かって直接話しかけ、D はその話を黙って聞いていた。手動ではこの時点で BD 間にイマジナリーラインを設定したが、提案手法は  $B \rightarrow D \rightarrow B$  と 3 ステップの発言を必要とするため、D と B の発言を待つ必要がある。実際、手動では BD 間に 7 秒設定されたが、提案手法では 5 秒遅れて設定されたため 2 秒間しか一致しなかった。2人の対話時にカバー率が低いのは、このような少ない発言数で設定される場合に対応が遅れる影響が大きい。

また、 $t_1$  後半ではしばらく 3人の議論が進行したが、うち 1人が途中から全員に向かって話を始めた。手動ではこの時点で解除されたが、提案手法は無関係な参加者の発言を基準にしているためこれを認識できず、解除が 13 秒遅れた。同様の例は  $t_3$  前半、 $t_5$  にも見られ、冗長部分が増えて有効率が低下した原因となっている。このように解除遅延は聞き手の変化が主な原因であるが、人数が 5 人だと議論中以外の参加者は 2人と少なく発言頻度が低いことも影響がある。逆に 2人の間のイマジナリーラインは、他の参加者が 3人いるため解除されやすく有効率が高くなったと考えられる。

この遅延は 2カ所で、編集映像の違和感につながっ

た。 $t_3$  前半では BCD の議論に表 3 のカメラ 2 (shot 1)、6 (shot 2)、8 (shot 2) が選択される (図 7 および表 3 参照)。その後、 $t_3$  中間部において手動では AC 間にイマジナリーラインが設定されたが、提案手法ではこれを認識できず、以前のイマジナリーラインが解除されなかった。先ほどの 3つのショットには A を映すものは 1 つもないため、A が重要な発言をしているにもかかわらず声だけが聞こえる映像となってしまった。同様の状況は  $t_6$  でも発生した。

しかしそれ以外の部分では、映像表現への影響は少なかった。 $t_1$ 、 $t_2$ 、 $t_4$  のように、より演出が効果を持つ長時間の対話・議論のカバー率は 80%を超えている。 $t_2$  では設定遅延が目立つが、これは  $A \rightarrow D \rightarrow A$  という流れの中で D の発言が長時間続いたことによる。この間、提案手法では D を含むショットと全景ショットの 2種類を切り替えるため、特に違和感のない映像となった。 $t_1$  後半や  $t_5$  で目立つ解除遅延では、手動でイマジナリーラインが解除された後に発言を続けたのが、直前までその設定に関与していた参加者であったため、 $t_2$  と同様に違和感のない映像を生成できた。

#### 5.4.2 カメラ配置の影響

映像全体を通しての平均得点は (a) 2.43、(b) 1.15、(c) 2.43 となり、(a) と (c) が高く評価された。

提案手法の配置である (a) は、開始から 3分までの B と D が積極的に発言した時間帯につねに評価が高かった。その理由として「表情がよく見える」というものがあつた。(a) の底辺カメラは BD 間のイマジナリーラインの近くに位置し、それぞれの主観的視点に近いショットが得られる。このため BD の対話時に「話しかけている」「話を聞いている」表情をより正面から撮影できたためだと考えられる。

一方、(b) は全体を通して評価が低く、「変化に乏しい」「横顔が多い」などのコメントがあつた。(b) の底辺カメラの位置は頂角カメラに近く、ショットも客観的なものになる。このため頂角カメラのショットとス

表 5 比較実験におけるアンケートの評価結果

Table 5 Results of the questionnaire data from the comparative experiment.

No	質問項目	人手	提案手法	Wilcoxon 符号付順位 検定 p 値
1	議論の流れがつかめた	2.81	3.94	**0.0017
2	だれとだれが会話しているかがよく分かった	2.69	4.06	**0.0017
3	映像に退屈しなかった	2.94	3.81	**0.0068
4	人物の位置関係がよくつかめた	3.13	3.88	*0.0107
5	カメラの切替に違和感を感じなかった	2.50	3.25	*0.0244
6	その場の状況が分かりやすかった	3.19	3.81	*0.0269
7	見やすい映像だった	3.00	3.44	0.1309
8	話し手がよく分かった	4.19	4.00	0.4375
9	画面上の人物の表情や身ぶりがよく分かった	3.69	3.50	0.4961
10	見たい映像に切り替わっていた	3.31	3.25	0.8438
11	画面上の人物の存在感があった	3.38	3.38	0.9999
12	切替えのタイミングは適切だった	3.06	3.06	0.9999

(N=16; \*\*:p &lt; 0.01, \*:p &lt; 0.05)

イチングを行っても、各ショットにあまり差異が見られず、平面的な映像になってしまうと考えられる。

(c) の得点は、D と C が B に対して積極的に発言し、B が聞き役にまわっていた時間帯（3～5分）で特に評価が高くなった。発言頻度の極端な偏りにより D が重要人物と認識され、頂角カメラも D の主観的視点に偏った構図が評価されたようだ。底辺の大きな三角形にカメラを配置する提案手法は、我々の想定した双方が発言を繰り返すシーンには適しているといえるが、発言頻度と配置に関して会議の種類や進行方法を含めて新たに検討してみる価値があると思われる。

#### 5.4.3 アンケート回答結果の分析

アンケートの結果を表 5 に示す。各質問は「まったくあてはまらない」、「あまりあてはまらない」、「どちらともいえない」、「ややあてはまる」、「かなりあてはまる」の 5 段階にそれぞれ 1 点から 5 点を与え、映像別に各質問に対する平均得点を求めた。さらに、評点に有意差があるか確認するため Wilcoxon の符号付順位検定 p 値を求めた。表中の人手・提案手法の各項目の値は評価値の平均得点である。

提案手法による映像はいくつかの項目で人手の映像を上回る評価を得ることができた。項目 1 および 2 の結果から、人手よりも議論や会話の様子が分かりやすかった。また項目 4, 5, 6 からは、参加者の位置関係が明確で違和感の少ない映像であるという評価も得た。提案手法によるイマジナリーラインの検出と撮影カメラの決定が演出効果として現れたのが分かる。

これに対し、項目 8, 9, 11 では人手の映像に及ばなかった。これらの項目はいずれも参加者単独の映り具合に関わるものである。今回用いたショットはすべて複数の参加者を映す構図であり、参加者全体の様子

の伝達には効果があったものの、個人の表情や存在感は薄らいってしまったといえる。画像処理などで被写体の位置を補正してアップショットを適宜実現するなど、ショットの種類を増やしていく必要がある。

同様に、ショットの持続時間に関する項目 10 および 12 においても人手の映像の方がわずかに高い評価を得た。本研究では、落ち着いた/白熱したといった会話内容を考慮しないため、長めの持続時間が望ましい落ち着いた状態においても、強制的に次のショットへ移行してしまうケースが見られた。人手による映像はカメラマンが会話内容を判断したため評価が高くなったものと考えられる。

また、アンケート項目以外のコメントには「映像が跳ねている」ように感じるという意見があった。図 8 の各ショットは B と C の画面上の位置がずれている。このため、これらショットを接続すると参加者が振動しているように見えてしまった。ショット接続の際に人物の視線や位置を一定に保つなど、新たな映像理論を取り入れることで改善できると考えられる。

## 6. む す び

対面会議シーンを複数カメラで自動撮影する場合、従来議論されてきたショットの持続時間以外にも撮影カメラ位置の選択が重要となる。そこで本論文では映画の撮影技法を利用した自動撮影手法を提案した。提案手法は参加者の発言の推移からイマジナリーラインを検出し、それを基準に三角形の配置になるよう撮影カメラを決定する。評価用映像では、提案方式のイマジナリーラインのカバー率は 70%、有効率は 61%であり、この検出ロスが映像の違和感につながる箇所は少なかった。カメラの配置方法は、参加者が発言を繰



り返すシーンにおいて効果があり、手で撮影カメラを選択した場合と比較して位置関係の分かりやすい映像を自動生成することができた。

一方、少ない発言数でのイマジナリーラインの設定と、聞き手が途中で変化した場合の解除には対応できず、冗長な部分が多いことも分かった。また、ショットの構図と持続時間に課題も残った。この改善には顔の向きや会話内容の判断といった、より高度な認識技術を取り入れる必要がある。これ以外にも、発言頻度によるカメラ配置の変動、適用人数の拡大や新たな映像理論を導入することで、より効果的な映像を自動撮影できるよう完成度を高めていきたいと考えている。

謝辞 本研究の一部は、21世紀COEプログラム研究拠点形成費補助金のもとに行われた。ここに記して謝意を表す。

### 参 考 文 献

- 1) 大西正輝, 泉 正夫, 福永邦雄: 情報発生量の分布に基づく遠隔講義撮影の自動化, 電子情報通信学会論文誌, Vol.J82-D-II, No.2, pp.1590-1597 (1999).
- 2) 先山卓朗, 大野直樹, 椋木雅之, 池田克夫: 遠隔講義における講義状況に応じた送信映像選択, 電子情報通信学会論文誌, Vol.J84-D-II, No.2, pp.248-257 (2001).
- 3) Ozeki, M., Nakamura, Y. and Ohta, Y.: Camerawork For Intelligent Video Production — Capturing Desktop Manipulations, *International Conference on Multimedia and Expo (ICME2001)*, pp.41-44 (2001).
- 4) Lee, D.-S., Erol, B., Graham, J., Hull, J.J. and Murata, N.: Portable meeting recorder, *Proc. 10th ACM International Conference on Multimedia*, Juan-les-Pins, France, pp.493-502 (2002).
- 5) Cutler, R., Rui, Y., Gupta, A., Cadiz, J., Tashev, I., wei He, L., Colburn, A., Zhang, Z., Liu, Z. and Silverberg, S.: Distributed meetings: a meeting capture and broadcasting system, *Proc. 10th ACM International Conference on Multimedia*, Juan-les-Pins, France, pp.503-512 (2002).
- 6) 井上智雄, 岡田謙一, 松下 温: テレビ番組のカメラワークの知識に基づいたTV会議システム, 情報処理学会論文誌, Vol.37, No.11, pp.2095-2104 (1996).
- 7) 大西正輝, 影林岳彦, 福永邦雄: 視聴覚情報の統合による会議映像の自動撮影, 電子情報通信学会論文誌, Vol.J85-D-II, No.3, pp.537-542 (2002).
- 8) 西山晴彦, 大久保達真, 斉藤伸介, 松下 温: 映像の知識に基づく仮想空間演出, 電子情報通信学会論文誌, Vol.J81-D-II, No.1, pp.146-155 (1998).
- 9) Christianson, D., Anderson, S., wei He, L., Salesin, D., Weld, D. and Cohen, M.: Declarative Camera Control for Automatic Cinematography, *Proc. AAAI '96*, pp.148-155 (1996).
- 10) 宮崎英明, 亀田能成, 美濃導彦: 複数のカメラを用いた複数ユーザに対する講義の実時間映像化手法, 電子情報通信学会論文誌, Vol.J82-D-II, No.10, pp.1598-1605 (1999).
- 11) 熊野雅仁, 有木康雄, 上原邦昭, 下条真司, 春藤憲司, 塚田清志: 映像編集支援システムのためのショットサイズ自動付与, 電子情報通信学会論文誌, Vol.J85-D-I, No.7, pp.592-602 (2002).
- 12) 平石絢子, 井上亮文, 重野 寛, 岡田謙一, 松下 温: 映画の撮影手法に基づいた会議の自動撮影, マルチメディア, 分散, 協調とモバイル (DICOMO2002) シンポジウム論文集, pp.285-288 (2002).
- 13) ダニエル・アリホン: 映画の文法 実作品に見る撮影と編集の技法, 紀伊国屋書店 (1999).
- 14) Drozd, A., Bowers, J., Benford, S., Greenhalgh, C. and Fraser, M.: Collaboratively Improving Magic: An Approach to Managing Participation in an On-Line Drama, *Proc. EC-SCW2001*, Bonn, Germany, pp.159-178 (2001).

(平成 15 年 5 月 27 日受付)  
(平成 15 年 11 月 4 日採録)



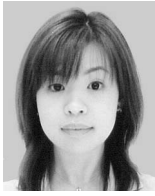
井上 亮文 (学生会員)

1999年慶應義塾大学理工学部計測工学科卒業。2001年同大学院前期博士課程修了。現在同大学院後期博士課程在学中。分散コンピューティング、自動撮影技術の研究に従事。



吉田 竜二

2000年慶應義塾大学理工学部情報工学科卒業。2002年同大学院修士課程修了。現在日本ヒューレット・パッカー株式会社勤務。在学中、自動撮影技術の研究に従事。



平石 絢子

2001年慶應義塾大学理工学部情報工学科卒業。2003年同大学院修士課程修了。現在株式会社NTTドコモに勤務。在学中、自動撮影技術の研究に従事。



重野 寛(正会員)

1990年慶應義塾大学理工学部計測工学科卒業。1997年同大学院理工学研究科博士課程修了。1998年同大学理工学部情報工学科助手(有期)。現在、同大学理工学部情報工学科助教授。博士(工学)。計算機ネットワーク・プロトコル、モバイル・コンピューティング、マルチメディア・アプリケーション等の研究に従事。情報処理学会マルチメディア通信と分散処理研究会幹事。著書「～ネットワーク・ユーザのための～無線LAN技術講座」(ソフト・リサーチ・センター)、「コンピュータネットワーク」(オーム社)等。電子情報通信学会、IEEE、ACM各会員。



岡田 謙一(フェロー)

慶應義塾大学理工学部情報工学科教授、工学博士。専門は、CSCW、グループウェア、コンピュータ・ヒューマン・インタラクション。「ヒューマンコンピュータインタラクション」(オーム社)、「コラボレーションとコミュニケーション」(共立出版)をはじめ著書多数。情報処理学会誌編集主査、論文誌編集主査、GW研究会主査等を歴任。現在、情報処理学会GN研究会運営委員、BCC研究グループ幹事、日本VR学会仮想都市研究会副委員長。IEEE、ACM、電子情報通信学会、人工知能学会各会員。1995年度情報処理学会論文賞、情報処理学会40周年記念論文賞、2000年度情報処理学会論文賞受賞。2002年情報処理学会フェロー。



松下 温(フェロー)

1963年慶應義塾大学工学部電気工学科卒業。1968年イリノイ大学大学院コンピュータサイエンス専攻修了。1989年4月より2002年3月まで慶應義塾大学理工学部教授。2002年4月より東京工科大学コンピュータサイエンス学部長、教授および慶應義塾大学理工学部客員教授。工学博士。マルチメディア通信、コンピュータネットワーク、グループウェア等の研究に従事。情報処理学会理事、同学会副会長、マルチメディア通信と分散処理研究会委員長、グループウェア研究会委員長、MIS研究会委員長、バーチャルリアリティ学会サイバースペースと仮想都市研究会委員長等を歴任。現在、情報処理学会ITS研究会主査。郵政省、通産省、建設省、農水省、都市基盤整備公団、行政情報システム研究所等の委員長、座長、委員を多数歴任。「やさしいLANの知識」(オーム社)、「201x年の世界」(共立出版)等著書多数。1993年情報処理学会ベストオーサ賞、1995年および2000年情報処理学会論文賞、2000年10月情報処理学会40周年記念90年代学会誌論文賞、2000年10月電子情報通信学会フェロー、2000年10月VR学会サイバースペース研究会、2001年5月情報処理学会功績賞。2002年3月情報処理学会フェロー、電子情報通信学会、人工知能学会、ファジィ学会、IEEE、ACM各会員。